Using dynamic geospatial ontologies to support information extraction from big Earth observation data sets

Gilberto Camara[1], Adeline Maciel[1], Victor Maus[1], Lubia Vinhas[1], and Alber Sanchez[1]

[1]Image Processing Division, National Institute for Space Research, Av dos Astronautas 1758, Sao Jose dos Campos, 12227-001 Brazil.
Email: {gilberto.camara, adeline.maciel, victor.maus, lubia.vinhas, alber.ipia}@inpe.br

Abstract

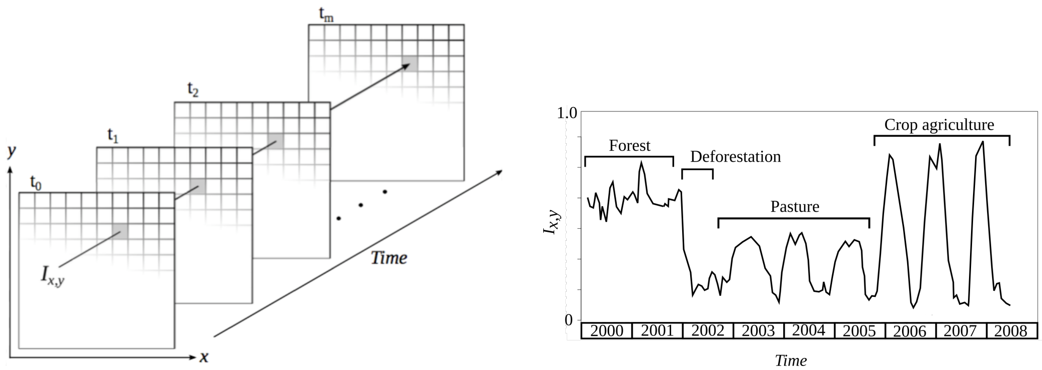
This paper presents the spatiotemporal interval logic formalism and shows how to use it for reasoning about land use change using big Earth observation data. This formalism improves our ability to extract information from large land remote sensing data sets.

1 Events as key concepts for describing land use change

Remote sensing satellites are the only source that provides consistent data about the Earth's land and oceans. The open availability of big Earth observation data has led to an opportunity to improve information on land changes in the planet. However, most studies that use remote sensing images to detect change still adopt a *snapshot* approach. Image from a sequence are classified one by one; results are compared to account for change. There is no actual representation of the occurrences of change, but only of their effects. Two land areas with different change trajectories whose initial and final states are the same cannot be distinguished. With access to big data sets, researchers need better ways to describe and understand change. The challenge is to make best use of big Earth observation data sets to represent change.

This paper uses the concept of *'events'* from dynamic spatial ontologies to describe land use change. Events are complete entities on their respective time intervals; their lifetime is limited while objects persist in time and are complete in space (Galton and Mizoguchi, 2009; Worboys, 2005; Hacker, 1982). Since events are intrinsically related to the objects they modify, a geospatial event calculus should specify not only what happens, but also which objects are affected by such changes. We present an event calculus formalism for reasoning about land use change. The formalism is general enough to be applied in other geospatial domains.

To define events in big Earth observation data sets, multiple satellite observations of an area are mapped to 3D arrays in space-time. A pixel location (x, y) in consecutive times $t_1, ..., t_m$ makes up a satellite image time series (Figure 1a). One can extract land use change information for each pixel, considered as an atomic 'land object'. Data mining techniques such as Time-Weighted Dynamic Time Warping (Maus et al., 2016) match temporal patterns of events to their actual occurrence in remote sensing time series (Figure 1b). The results are the temporal boundaries of events associated to a land object. For example, Figure 1b shows four major events extracted from a remote sensing time series, expressed in terms of the intervals they happen. From 2000 to 2001 the area was a forest that was deforested in 2002. From 2003 to 2005 the area it was used for pasture and from 2005 to 2008, as cropland. Since classifying all pixels in a space-time array produces a large set of events, we need an event reasoning formalism to extract information.



(a) Time series measures (e.g. EVI index) of a pixel location (x, y)

(b) Land-use/cover change events associated to a pixel location (x, y)

Figure 1: A 3-dimensional array of satellite data and events describing change at a particular location. Adapted from Maus et al. (2016).

2 The spatiotemporal interval logic

The main elements of a temporal reasoning formalism include the primitive time unit (*instants or intervals?*) and the granularity (*is time continuous or a sequence of discrete elements?*). For describing land use change trajectories from remote sensing data, we consider that an interval-based approach with discrete granularity is better than instant-based formalisms such as the Event Calculus (Kowalski and Sergot, 1989). Thus, we propose to extend Allen's interval temporal logic (Allen, 1984) to build a general framework to reason about events. Allen (1983) defines a set of mutually exclusive primitive relations between temporal intervals. Each of these is a predicate over intervals: *during*, *starts*, *finishes*, *before*, *overlap*, *meets*, and *equal*. These predicates have become widely used in many areas of computing.

In this work, we propose a spatiotemporal interval logic that includes geospatial objects explicitly. Geo-objects are intrinsically tied to space, and events change their properties. The elements of the formalism are a set of *discrete geo-objects* ($O = o_1, o_2, ..., o_n$), *discrete time intervals* ($T = t_1, t_2, ..., t_n$), and *properties of objects* ($P = p_1, p_2, ..., p_n$). Extending the ideas from Allen (1984), we introduce the predicate $holds(o, p, t) \rightarrow$ *bool*, to denote the assertion that the property p of geo-object o holds over interval t. We also introduce the predicate $occur(o, p, T_e) \rightarrow bool$ to denote that, given an interval $T_e \subset T$, the property p of geo-object o is true over the whole subset T_e. Some of the basic axioms of our spatiotemporal interval logic are presented in Table 1. In these axioms, we use the notation $T_e \subset T$ to denote a temporally connected proper subset of T.

Table 1: Basic axioms of spatiotemporal interval logic

Events happen over a given interval $\forall o \in O, occur(o, p, T_e) \wedge in(T'_e, T_e) \implies \neg occur(o, p, T'_e)$ *where* $in(T'_e, T_e) \Leftrightarrow during(T'_e, T_e) \vee starts(T'_e, T_e) \vee finishes(T'_e, T_e)$
Events do not change over an interval $\forall o \in O, occur(o, p, T_e) \implies \forall t \in T_e, holds(o, p, t)$
Events are unique $\forall o \in O, occur(o, p, T_e) \wedge meets(T_e, T'_e) \wedge occur(o, p', T'_e) \implies p \neq p'$

3 Reasoning about land use change

While the full development of the spatiotemporal interval logical applied to land classification is beyond the scope of this paper, we show some queries useful to reason about *land use trajectories*. Informally, a *land use trajectory* is a path from one land use state to another, for example when a forest area is converted to pasture. Formally land use trajectories are expressed as logical expressions over an event data set.

As an example, consider a study that investigates the agreement known as the Brazil's Soy Moratorium, signed by major commodity traders agreeing not to purchase soybeans grown on lands deforested after July 2006 in the Brazilian Amazonia (Gibbs et al. (2015)). Farmers abiding by the Soy Moratorium agree not to directly replace forest by soybean plantations. However, the agreement does not preclude indirect land use changes, as when a farmer buys land previously deforested that is being used as pasture. In this case, the cattle rancher may sell his land and move elsewhere, causing deforestation without violating the Soy Moratorium. Thus, we want to discover not only direct land use changes, where forest is replaced by soybeans, but also indirect land use changes. The queries in Table 2 point out how to elicit both direct and indirect land use change caused by soybeans in Amazonia.

Table 2: Using the spatiotemporal interval logic to map land use change trajectories in Brazil

Which forest areas have been replaced by soybeans? $occur(o, \text{"}forest\text{"}, t_1) \wedge meets(t_1, t_2) \wedge occur(o, \text{"}deforestation\text{"}, t_2)$ $\wedge\, meets(t_2, t_3) \wedge occur(o, \text{"}soy\text{"}, t_3)$
Which forest areas have been replaced by pasture and later turned into soybean? $occur(o, \text{"}forest\text{"}, t_1) \wedge meets(t_1, t_2) \wedge occur(o, \text{"}deforestation\text{"}, t_2) \wedge meets(t_2, t_3)$ $\wedge\, occur(o, \text{"}pasture\text{"}, t_3) \wedge meets(t_3, t_4) \wedge occur(o, \text{"}soy\text{"}, t_4)$

4 Conclusions

This paper presents the spatiotemporal interval logic, which is a spatial extensions of the temporal interval logic proposed by Allen (1984). The formalism considers the nature of events detectable using Earth observation data, which are discrete transitions where one land cover type is replaced by another. The proposed logic allows reasoning about land use trajectories in regional and global areas. To be useful, this formalism needs to be supported by efficient data mining techniques, capable of extracting event data sets from big data. When such event data be available, the spatiotemporal interval logic improves information extraction from large remote sensing data sets.

Acknowledgements

The authors are supported by FAPESP e-science program (grants 2014-08398-6 and 2016-03397-7). Gilberto Camara is also supported by CNPq (grant 3121512014-4).

References

J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.

J. F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154, 1984.

A. Galton and R. Mizoguchi. The water falls but the waterfall does not fall: New perspectives on objects , processes and events. *Applied Ontology*, 4:71–107, 2009.

H. K. Gibbs, L. Rausch, J. Munger, et al. Brazil's soy moratorium. *Science*, 347(6220):377–378, 2015.

P. Hacker. Events and Objects in Space and Time. *Mind*, XCI(361):1–19, 1982.

R. Kowalski and M. Sergot. A logic-based calculus of events. In *Foundations of knowledge based management*, pages 23–55. Springer, 1989.

V. Maus, G. Câmara, R. Cartaxo, A. Sanchez, F. M. Ramos, and G. R. de Queiroz. A time-weighted dynamic time warping method for land-use and land-cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, PP(99):1–11, 2016.

M. Worboys. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science*, 19:1–28, 2005.